Jour. Ind. Soc. Ag. Statistics
55(3), 2002 : 297-322

A Study on AMMI Model and its Biplots[1]

B.M.K. Raju
Indian Agricultural Statistics Research Institute, New Delhi

SUMMARY

AMMI model has been shown to be a useful technique to capture the non-linear interactions, when Joint Regression technique fails to perceive important effects in studies of G × E interaction. The application of biplots to draw reliable stability conclusions is subject of great interest when significant proportion of interaction explained is by the first or first two PCA axes. In the present study some stability measures are proposed which are equivalent to biplot with first PCA axis and biplot with first two PCA axes for ranking purposes. The reliability of stability conclusions improves with increase in the number of PCA axes, which has been exploited while proposing the new measure of stability. The proposed stability measure $W_{i(AMMI)}$ which accommodates all possible PCA axes is shown equivalent to Wricke's ecovalence. The proposed stability measures are precise in the order in which amount of information increases. The ranking ability of $W_{i(AMMI)}$ is found to be superior to other measures when there are missing cells in the data; showing some kind of robustness to the missing data. The ranking abilities of different stability measures are found to be better in the proposed EM-AMMI with random environments as compared to EM-AMMI of Gauch and Zobel [6] and Modified EM-AMMI of Bajpai [1] revealing its superiority over the other two methodologies. Thus, the stability measure $W_{i(AMMI)}$ using EM-AMMI with random environments methodology may be recommended and may be employed to derive stability conclusions from AMMI model when some cells in two-way table are missing. EM-AMMI enriched technique has been proposed for Joint Regression to deal with incomplete data. This technique enables us to fit not only the main effects but also the interactions for the missing cells.

Key words : Joint regression, Predictive success, Stein effect, Stability, Biplot, Wricke's ecovalence, Incomplete data, EM-AMMI, Indirect data.

[1] Part of the Ph.D. research work done under the guidance of Dr. V.K. Bhatia, Principal Scientist, IASRI and the paper was presented in the session on ISAS Young Scientist Award during the 56th Annual Conference of the Indian Society of Agricultural Statistics held at UAS, Dhawad from 18th to 20th December, 2002 and adjudged best for ISAS Young Scientist Award 2002.

1. Introduction

Recently yielding stability as a selection trait in plant breeding programmes and evaluation trials is constantly gaining importance over yielding ability especially in developing countries like India, where the number of small and marginal farmers with small holdings is very high. A main strategy among small-scale subsistence farmers, particularly in marginal areas, is risk minimization. In these areas, stable yields are the key to sustainable food supplies. However the existence of interaction reflecting differences among varieties in their ability to maintain performance over a wide range of environmental conditions is long been recognized to exist. Hence the challenge put forward for the plant breeder has been to develop cultivars that are stable across a range of environments. If one follows the dynamic concept, the goal of breeding stable genotypes may be translated as the goal of minimizing genotype environment interaction, which makes the selection of high yielding genotypes easier.

As an alternative to additive ANOVA model, which identifies the interaction as a source but does not analyze it, multiplicative formulations may be chosen to quantify the variety's contribution to genotype × environment interaction, which include well known Joint Regression and at the moment most popular Additive Main effects and Multiplicative Interaction (AMMI) model. These multiplicative formulations permit the interpretation of interaction as differential genotypic sensitivity to environmental variable(s).

Joint Regression consists in regressing the observed yield on the observed environmental mean yield. The resulting regression coefficient may be interpreted as linear sensitivity of the variety to environmental change. However this procedure suffers from a conceptual problem of regressing a vector of observations on another vector, which is a linear combination of the former. Hence the estimates of sensitivity obtained from this method are biased. Further, when the component of deviation from linear regression is significant, Joint Regression fails to perceive important interaction effects. This makes the prediction of variety's behaviour over a range of environments imprecise. Subsequently, one may not be able to make any remarks on the stability. In such cases, Additive Main effects and Multiplicative Interaction (AMMI) model proves to be a more realistic one in the sense that it can digest the non-linear interactions too into a pattern rich model, discarding a noise rich residual. This methodology consists in decomposing the matrix of residuals, obtained after fitting the additive main effects, by singular value decomposition. Graphical representation of interaction using AMMI interaction parameters is known as biplot. Till date, the stability conclusions made from AMMI model are based on biplots. However the scope of biplots is very much limited. Biplot formulation of interaction will be successful only when significant proportion of G × E interaction is concentrated in the first or first two PCA axes. Thus in view of

this, one may be interested in deriving a more comprehensive stability measure from AMMI model.

Most often the data obtained from multi environment testing are unbalanced. It may be incidental or accidental. There is a need to study the existing methodologies for such data sets also and to have further modification for their improvement. Further, treating the environment effect as random will have some desirable consequences in the sense that the corresponding BLUP estimator corrects for possible random environmental contributions to the genetic effects. Due to this, the bias creeping in due to the selection of environments is precluded. The aim of this study is to determine and develop a reliable and statistically sound methodology, that efficiently explains the genotype environment interaction, and a practical solution of obtaining reliable and stable genotypes, keeping in view of the said issues and concerns.

2. Joint Regression for Balanced Data

Eberhart and Russell [2] proposed an observational formulation for the Joint Regression context. The model proposed by Eberhart and Russell [2] may be written as

$$y_{ij} = \alpha_i + \beta_i e_j + \delta_{ij}$$

where

y_{ij} is the performance of i-th genotype at the j-th environment ($i = 1, \ldots, K$; $j = 1, \ldots, N$) averaged over R replications

α_i is the mean of i-th genotype over all the environments

e_j is the environmental index for the j-th environment which may be obtained as the mean of all genotypes at the j-th environment minus the general mean

β_i is the regression coefficient measuring the linear sensitivity of i-th genotype to environment change

δ_{ij} is the 'deviation from regression' of the i-th genotype in the j-th environment

The estimates of parameters may be obtained as following:

$$\hat{\alpha}_i = \bar{y}_{i.}$$

$$\hat{e}_j = \bar{y}_{.j} - \bar{y}_{..} \text{ such that } \sum_j e_j = 0$$

$$\hat{\beta}_i = \frac{\sum_j y_{ij} e_j}{\sum_j e_j^2}$$

Mean squared deviation from linear regression, $\overline{S}^2_{d_i}$ may be obtained as

$$\overline{S}^2_{d_i} = \sum_j \frac{\delta^2_{ij}}{N-2} - \frac{S^2_e}{R}$$

Stability and Adaptability

A genotype with unit regression coefficient i.e. $\beta_i = 1$ and the mean squared deviation not significantly different from zero ($\overline{S}^2_{d_i} = 0$) is said to be stable. Significance of $\overline{S}^2_{d_i}$ from zero invalidates the linear prediction. If $\overline{S}^2_{d_i}$ is not significantly different from zero, the performance of the genotype for a given environment may be predicted. Accordingly, a genotype whose performance can be predicted is said to be stable and it also helps in choosing genotypes for specific adaptation.

3. AMMI Model and its Evaluation for a Balanced Data

The AMMI model for a two way table of genotype × environments may be written as

$$y_{ij} = \mu + \alpha_i + \beta_j + \sum_{m=1}^{m'} \lambda_m \gamma_{mi} \delta_{mj} + \theta_{ij},\ i = 1, \ldots, K \text{ and } j = 1, \ldots, N \quad (1)$$

where

y_{ij} is the mean yield of i-th genotype in j-th environment

μ is the grand mean

α_i is the i-th genotype mean deviation

β_j is the j-th environment mean deviation

m' is the number of PCA axes retained in the model

λ_m is the singular value for the PCA axis, m

γ_{mi} is the i-th genotype PCA score for the axis, m

δ_{mj} is the j-th environment PCA score for the axis, m

θ_{ij} is the residual

The identification constraints for the model (1) are as under

$$\sigma^2_i \ \forall\ m \quad (2)$$

$$\sum_i \gamma^2_{mi} = 1 = \sum_j \delta^2_{mj},\ \forall\ m \quad (3)$$

$$\sum_i \gamma_{mi}\gamma_{m^*i} = 0 \text{ and } \sum_j \delta_{mj}\delta_{m^*j} = 0 \text{ where } m \neq m^* \quad (4)$$

Ordinarily the number m' of interaction principal component axes retained in the model is chosen with empirical considerations of F test of significance, predictive accuracy, agricultural interpretability of the associated interaction PCA scores, and so on. The residual combines the M-m' discarded axes, where M = min [(K-1), (N-1)]. Equations (2) and (3) state that the vectors γ_{mi} and δ_{mj} are normalized. According to equation (4), the vectors γ_{mi} and γ_{m^*i} are orthogonal with a similar statement for δ_{mj} and δ_{m^*j}.

One way of constructing AMMI is that the interaction is described in terms of differential sensitivity to the most discriminating environmental variables that can be constructed. These environmental variables are hypothetical and obtained from the data themselves. No explicitly measured environmental variables enter the model. Because both environmental variables and genotypic sensitivities are estimated from the data table itself, the AMMI model is called a bilinear model: given the column parameters the model is linear in the row parameters and given the row parameters the model is linear in the column parameters. The basic model is essentially a two way ANOVA model, which requires that the matrix of interaction parameters be decomposed by using factor analytic techniques.

Let us reparameterize the equation (1) so as to obtain the matrix of interaction parameters as

$$y_{ij} = \mu + \alpha_i + \beta_j + V_{ij} \tag{5}$$

where $V_{ij} = \sum_{m=1}^{m'} \lambda_m \gamma_{mi} \delta_{mj} + \theta_{ij}$ of equation (1)

Now the estimates of V_{ij} may be obtained as

$$\hat{V}_{ij} = y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$$

Form the matrix X of interaction estimates from $\hat{V}_{ij}$'s such that each row of X denotes the interactions of a variety over N environments. Using factor analytic decomposition, the matrix X may be written as

$$X = ADB' \tag{6}$$

where

X is K × N matrix with $\hat{V}_{ij}$'s as elements

A is K × M orthonormal matrix

D is M × M diagonal matrix with elements $d_1 \geq d_2 \geq \ldots \geq d_{m'} \geq \ldots d_M$

B is N × M orthonormal matrix

M is the rank of X

The matrices A, D and B of equation (6) may be obtained from the characteristic vectors and characteristic roots of the K × K matrix XX'. The K × M matrix A then consists of the characteristic vectors and the M × M

diagonal matrix D consists of the square roots of the characteristic roots of XX'. The N × M matrix B can then obtained by solving

$$B = X'AD^{-1} \tag{7}$$

The above solution specifies that the matrices D and A be found by solving the eigenvalues and eigen vectors of the matrix XX' and then the matrix B be obtained from (7). It is also possible to solve for the matrices D and B by finding the eigenvalues and eigen vectors of the matrix X'X and then obtaining A from $A = X BD^{-1}$. For ease of calculation it is convenient to solve for the eigenvalues and eigen vectors of either of X'X or XX' whichever has the smaller dimension.

The environmental eigen vector corresponding to λ_1 (first column of B) represents the hypothetical environmental variable that describes the largest amount of interaction and thus best discriminates between genotypes, the second axis the second largest amount, and so on. If all the M possible axes are retained in the model, it completely factors out the interaction without leaving any residual. Multiplicative modelling of interaction is successful when the additive ANOVA interaction with (K – 1)(N – 1) independent parameters, can be replaced by only a few multiplicative terms (m' <<M), thus adequately describing the interactions with considerably fewer parameters. Various methods exist to determine the number of multiplicative terms that should be retained. The most simple one is due to Gollob [7], and suffices for many practical applications (Gauch [4]). For each axis a mean square is calculated, that is compared with the estimate of residual by means of an F test. The mean squares are obtained as follows:

The sum of squares for axis m is equal to the square of the singular value, λ^2_m. The corresponding number of degree of freedom is K + N – 1 – 2m. The required mean square is the quotient of these two quantities.

4. Predictive and Postdictive Success

The ability of a model to predict validation data, not used in constructing the model constitutes predictive success.The ability of reduced, parsimonious model to fit selfsame data (the data used in constructing the model) constitutes postdictive success, for example the percentage of treatment SS accounted for by a reduced model. Gauch [3] recommends Gollob's [7] simple F test as a practical approach to the evaluation of postdictive success.

Evaluating the Predictive Accuracy

In prediction, one data set is used to construct a model, while different and independent data are used to validate the model. For example an AMMI model can be fitted to some yield data, and its expected values can then be evaluated by calculating the root mean square prediction difference between model and validation observations not used previously in modeling. This use of independent validation data precludes bias. The root mean square predictive

difference (RMSPD) between AMMI model and validation observations is simply the square root of the quantity of the sum of squared differences between AMMI estimates and validation observations, divided by the number of validation observations (Gauch and Zobel [5]). RMSPD is in the same units as the yield measurements and a small value indicates predictive success or accuracy.

A model's predictive accuracy may be estimated as follows: consider the variance of a model σ^2_M, the variance of validation observations, σ^2_V and the variance of differences between model and validation observations, σ^2_{MV}. By the variance rule, $\sigma^2_{MV} = \sigma^2_M + \sigma^2_V$. Now, σ^2_{MV} can be estimated empirically as the mean square difference between the model's estimates and validation observations (i.e. as the square of RMSPD). Likewise σ^2_V is estimated empirically by the Error MS. Thus the model's accuracy may be assessed by $\sigma^2_M = \sigma^2_{MV} - \sigma^2_V$. This estimate is unbiased because σ^2_{MV} and σ^2_V are both unbiased. In otherwords σ^2_M really does assess accuracy, not merely precision.

Furthermore, the model's predictive accuracy can be expressed in terms of number of effective replications, namely, the Error MS divided by σ^2_M. When the effective replications exceed the actual replications supplied to a model, the model exhibits the Stein effect, which implies the model is predictively accurate and the model is better than its data.

5. Biplots

Graphical display of interaction with AMMI interaction parameters is known as Biplot.

Let us distribute the singular values, λ_m over the genotypic scores, $\gamma^*_{mi} = \gamma_{mi} \lambda_m^c$ and the environmental scores $\delta^*_{mj} = \delta_{mj} \lambda_m^{(1-c)}$; where c is a scaling constant, varies from 0 to 1. The features of the biplot, however are not too critically dependent on c, and c = 0.5 may suit well for most problems. Two kinds of plotting is possible with estimated AMMI interaction parameters.

5.1 Biplot with First PCA Axis

First PCA scores of genotypes and environments are plotted against their respective means. This biplot formulates the interactions as $E(X_{ij}) = \lambda_1 \gamma_{1i} \delta_{1j} = \gamma^*_{1i} \delta^*_{1j}$, where X_{ij} is the interaction of i-th genotype in the j-th environment. Now the pattern of G × E interaction may be visualized from this plot. If the genotype or an environment has a PCA score of nearly zero, it will have smaller interaction effects. If a genotype and an environment are having the same sign on the PCA axis, their interaction is positive, if different, their interaction is negative. This biplot may also be used in simultaneous selection for yield and stability, provided the first PCA axis explains significant proportion of interaction SS.

5.2 Biplot with First Two PCA Axes

For a better description of the interaction, both first and second PCA scores of genotypes and environments may be considered for plotting. Here second PCA scores of genotypes and environments are plotted against their respective first PCA scores. The interaction from this biplot may formulated as $E(X_{ij}) = \gamma^*_{1i}\delta^*_{1j} + \gamma^*_{2i}\delta^*_{2j}$. The scores determine the end points of genotypic and environmental vectors, which depart from the origin. Simple geometry reveals that the interaction between a genotype i and an environment j can be obtained from a projection of either vector on to the other. In any quadrant the interaction between a genotype and an environment will be positive.

The stability of a variety or an environment is determined by the end point of its vector from the origin (0, 0). The vector, which is nearer to the origin, will have lesser interaction effects, hence may be regarded as a stable one. Obviously the stability conclusion made based on this biplot will be more precise than the former one; since this biplot takes into consideration the second PCA axis too for explaining the interaction.

However the scope of biplots is very much limited. Biplot formulation of interaction will be successful only when significant proportion of G × E interaction is concentrated in the first or first two PCA axes. When postdiction (F tests) mandates to retain more than 2 axes in the AMMI model, the biplot formulation of interaction will fail. Subsequently, the stability conclusions made based on biplots will be imprecise. In such cases, one may be interested to consider the procedures that accumulate considerable proportion of interaction SS (all the m' significant PCA axes or all possible M axes) to make the stability conclusions more reliable. Another remark that can be made on biplots is that, when the stability differences among the genotypes are very close it will be difficult to distinguish the genotypes with respect to the stability.

6. Some Measures of Stability from AMMI Model

Keeping in mind, the limitations of biplots concerning stability conclusions, we may attempt to derive a more comprehensive stability measure, retaining all possible 'M' PCA axes.

Let us consider the matrix X of interaction residuals obtained from ANOVA, $X = [(V_{ij})]$ where V_{ij} is the interaction residual of i-th genotype in j-th environment, i = 1, ..., K and j = 1, . . . , N.

Obtain the square, symmetric matrix XX' of order K × K from X.

Obtain the positive eigenvalues $\lambda_1^2, \lambda_2^2, ..., \lambda_m^2, ..., \lambda_M^2$ of XX'; where M = rank(XX')

Obtain the eigen vectors $\gamma_1, \gamma_2, \ldots, \gamma_m, \ldots, \gamma_M$ corresponding to the eigenvalues $\lambda_1^2, \lambda_2^2, \ldots, \lambda_m^2, \ldots, \lambda_M^2$ of XX', where γ_m is K × 1 vector, contains the PCA scores for the K genotypes corresponding to the axis m.

Using spectral decomposition, the square, symmetric matrix XX' may be written as

$$XX' = \lambda_1^2\gamma_1\gamma_1' + \lambda_2^2\gamma_2\gamma_2' + \ldots + \lambda_m^2\gamma_m\gamma_m' + \ldots + \lambda_M^2\gamma_M\gamma_M' \tag{8}$$

Let us consider the i-th diagonal element of XX' $\left(= \sum_{j=1}^{N} V_{ij}^2 \right)$, measures the spread of interaction effects of i-th genotype over N environments, which may be decomposed from (8) as

$$\begin{aligned}\sum_{j=1}^{N} V_{ij}^2 &= \lambda_1^2\gamma_{1i}\gamma_{1i}' + \lambda_2^2\gamma_{2i}\gamma_{2i}' + \ldots + \lambda_m^2\gamma_{mi}\gamma_{mi}' + \ldots + \lambda_M^2\gamma_{Mi}\gamma_{Mi}' \\ &= \lambda_1^2\gamma_{1i}^2 + \lambda_2^2\gamma_{2i}^2 + \ldots + \lambda_m^2\gamma_{mi}^2 + \ldots + \lambda_M^2\gamma_{Mi}^2 \\ &= \sum_{m=1}^{M} \lambda_m^2\gamma_{mi}^2\end{aligned}$$

However $\sum_{j=1}^{N} V_{ij}^2$ is a measure of stability proposed by Wricke [11], popularly known as Wricke's ecovalence. Hence the proposed measure of stability may be viewed as Wricke's ecovalence (W_i) in terms of AMMI parameters and may be denoted by $W_{i(AMMI)}$.

$$W_{i(AMMI)} = \sum_{m=1}^{M} \lambda_m^2\gamma_{mi}^2 \tag{9}$$

Therefore it may be concluded that the stability rank order obtained from the proposed measure will be equivalent to that of Wricke's ecovalence.

We may also develop few more measures of stability by retaining varying number of axes in the AMMI model.

When the first PCA axis only is retained in the AMMI model, then, we may measure the stability from FP_i as

$$FP_i = \lambda_1^2\gamma_{1i}^2 \tag{10}$$

However the absolute value of γ_{1i} alone is sufficient for comparison; since λ_1^2 is same for all the genotypes. Lesser the absolute value of γ_{mi}, more will be the stability. The comparison of genotypes for stability based on this measure will be equivalent to the comparison based on Biplot with first PCA axis.

We may retain the first two PCA axes in the AMMI model, to develop a measure B_i as

$$B_i = \sum_{m=1}^{2} \lambda_m^2 \gamma_{mi}^2 \tag{11}$$

Stability comparisons based on this measure will be equivalent to the comparisons based on Biplot with first two PCA axes.

We may also consider the measure based on fitted AMMI model by retaining m' axes, where m' is determined by the postdiction (F tests).

$$FA_i = \sum_{m=1}^{m'} \lambda_m^2 \gamma_{mi}^2 \tag{12}$$

In comparison to $W_{i(AMMI)}$, the above three measures will be less precise, as is evident from the fact that, they could not exploit the complete information. The reliability of a measure improves with the increase in the number of axes retained.

It may also interesting to propose, quantification for the accuracy of stability conclusions made based on FP_i, B_i and FA_i. If we assume that the stability rank order obtained from $W_{i(AMMI)}$ is true, then we may compare the remaining three measures of stability with respect to their ability to assess the true stability rank order. The concordance between the true rank order and the rank order displayed by the stability measure under consideration may be quantified by Spearman's rank correlation coefficient. Now the three measures FP_i, B_i and FA_i may be compared with respect to their estimated concordances.

For ranking purposes, the proposed measure $W_{i(AMMI)}$ will be equivalent to Shukla's unbiased estimator of stability variance, σ_i^2 (Shukla [9]).

7. AMMI for Incomplete Data

7.1 EM-AMMI

Implementation of AMMI requires that the two way table of interactions be complete. But, the interaction for the missing cells in incomplete Genotype × Environment table is undefined. Gauch and Zobel [6] suggested to employ Expectation and Maximisation (EM) algorithm to implement AMMI model for an incomplete two way table. They termed the so called missing data version of AMMI as "EM-AMMI". "In many important cases", including the AMMI model, "the EM algorithm is remarkably simple, both conceptually and computationally" (Little and Rubin [8]). In essence, EM involves "Filling in missing values and iterating" in such a manner that starting values do not affect the solution and hence are arbitrary and inconsequential, apart from some affect upon the number of iterations required for convergence. Little and Rubin [8]

summarize the computations: "the EM algorithm formalizes a relatively old ad hoc idea for handling missing data:

(1) replace the missing values by the estimated values

(2) estimate the parameters

(3) re-estimate the missing values assuming the new parameter estimates are correct

(4) re-estimate the parameters

and so forth, iterating until convergence."

A suitable implementation of the EM algorithm for EM-AMMI works as follows. First, compute cell means for every cell with data. Then initialize EM-AMMI's additive parameters by computing the un-weighted genotype means, environment means, and grand mean. Then initialize the interaction residuals as usual for cells with data (namely, the interaction equals the cell mean minus the genotype mean minus the environment mean plus the grand mean), but impute an interactions residual of zero for missing cells. Now, the interaction matrix has no unspecified cells, so perfectly ordinary PCA calculations solve for EM-AMMI's multiplicative parameters, continuing for as many PCA axes as desired. Next, re-estimate and revise each missing cell with the current EM-AMMI model. Then fit EM-AMMI to these revised data, treating imputed values the same as actual data. Iterate this process until convergence, i.e., until the imputed values for missing cells show acceptably small changes. Upon convergence, the EM-AMMI model "fits" the imputed cells perfectly with a residual of zero (within numerical precision), where as actual data have finite residuals as usual. Hence the EM algorithm fits a model to the actual data, while ignoring missing cells in the sense that they receive imputed values that fit the model perfectly.

Predictive Accuracy of Indirect Data

What we have described earlier under the head of predictive success or accuracy, is concerned to total data. To better understand, what the indirect data is, let us split the yield trial data consisting of K genotypes, N environments and R replicates into model data and validation data such that for each treatment (a cell in the two way table) R* (< R) replicates are chosen at random for AMMI modelling, while the remaining (R – R*) replicates are reserved for validation. This means that the model data consist of KNR* observations, where as the validation data consist of (KNR – KNR*) observations. Now the KNR* observation used to fit the AMMI model constitute the total (model) data; and for each treatment (i-th genotype grown in j-th environment) the remaining (KNR* – R*) modelling observations except its direct R* replications constitute the indirect information (data).

Estimation of yield for each treatment from indirect data is possible given an implementation of EM-AMMI described previously. For each

treatment in turn, EM-AMMI is given for only the indirect modelling data namely, the (KNR* – R*) observations, and the model is used to impute a yield value for the missing cell. The result is a complete matrix of imputed yield values for all KN treatments, in which not a single computation has used a single direct yield observation. We may now compute the Root Mean Square Predictive Difference (RMSPD) between these imputed values and the validation observations. Now, the predictive success of indirect information may be compared with the predictive success of total data in terms of their RMSPD values. Gauch and Zobel [6] illustrated this concept on a soyabean yield trial data. The amount of indirect EM-AMMI derived information (with an RMSPD of 395.4022) is found very nearer to the total information (with an RMSPD of 361.0331).

7.2 Modified EM-AMMI Approach

Bajpai [1] proposed to implement AMMI for an incomplete two-way data by initializing the EM-AMMI's additive parameters by Patterson's fitting constants technique. Here the additive main effects of equation (5) are estimated with "Fitcon" technique and then the interaction residuals are obtained from

$$V_{ij} = y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$$

Now proceed to impute the missing cells as in EM-AMMI.

Finally y_{ij}'s are estimated as

$$\hat{y}_{ij} = \hat{y}_{ij(fit)} + \hat{V}_{ij(AMMI)}$$

7.3 EM-AMMI with Random Environments

We may propose that when the environmental effects are treated as random, one may initialize the EM-AMMI's additive parameters with the ones obtained from mixed model equations. The mixed model equations provide the BLUE's for grand mean and genotypic effects and BLUP's for environmental effects. After fitting the additive main effects, the interaction residuals may be obtained as

$$V_{ij} = y_{ij} - \hat{\mu} - \hat{\alpha}_i - \tilde{\beta}_j$$

Now proceed to impute the missing cells as in EM-AMMI.

Comparison of the Three EM-AMMI Approaches

We may propose to compare the EM-AMMI approaches outlined here, to implement AMMI for an incomplete data, with respect to their ability to yield and stability conclusions as obtained from the complete data set using a similar approach. One may use the four measures of stability proposed with AMMI

model for this purpose. Assuming the stability rank order obtained from balanced data as true, we may quantify the concordance of stability rank order displayed by the methods of interest, that deal with incomplete data, by means of Spearman's rank correlation.

8. EM-AMMI Enriched Technique for Joint Regression for Unbalanced Data

Till date, while fitting the Joint Regression for incomplete data, for the missing cells, one has to compromise with main effects; since the interaction for a missing cell stays undefined. Appreciating the predictive success of indirect data exhibited by soyabean data of Gauch and Zobel [6], we may, now, propose an EM-AMMI enriched technique for Joint Regression for unbalanced data. EM-AMMI enables us to fit the interactions too for the missing cells. Having impution of the missing cells in the two way table of yields, with the ones obtained from an appropriate EM-AMMI, such that the two way table of yields is complete, evaluation of Joint Regression becomes straightforward.

8.1 EM-AMMI Substituted Least Squares for Joint Regression

Form the complete two-way table of yields by imputing the missing cells with modified EM-AMMI (Bajpai [1]) the estimates $\hat{e}_j$'s from the table can be obtained. Now one may regress the y_{ij}'s for each variety on the $\hat{e}_j$'s to evaluate the linear sensitivities.

8.2 EM-AMMI Substituted BLUP for Joint Regression

When the environmental effect is treated as random, form the complete two-way table of yields by predicting the missing cells with EM-AMMI with random environments and compute $\tilde{e}_j$. Now regress the y_{ij}'s for each i, on the $\tilde{e}_j$'s so obtained.

9. Empirical Study

Extent of Data

The data used in this study were collected from the multi-location year trials of released and pre-release varieties of groundnut conducted at research stations situated in different agro-climatic zones of Andhra Pradesh. The data were supplied by Regional Agricultural Research Station (RARS), Palem. The data consist of 20 environments and 15 genotypes. The experiments were laid in Randomized Block Design (RBD) with 3 replications. The pod yields were expressed as kg/ha. The mean data over the replicates for the 15 genotypes and 20 environments are presented in Table 1.

Joint Regression for Balanced Data

The genotypic parameters of Eberhart and Russell [2] model are presented in Table 2 and the corresponding analysis of variance is presented in Table 3. ANOVA in Table 3 reveals that there is significant difference among the genotypes indicating wider genetic diversity among the genotypes. Genotype x Environment (linear) and pooled deviation were found to be significant when tested against pooled error, indicate significant genotype x environment interaction. Genotype x Environment (linear) interaction was found to be not significant when tested against pooled deviation which implies that the genotypes don't differ for their regression on environmental index and overwhelming portion of G x E interaction is of non-linear type, which ultimately makes the behaviour of genotypes unpredictable. If we look at the significance of deviation from linear regression for the 15 genotypes in Table 3, all the deviations are significant at 1% level except genotypes G-7 and G-14. The deviation for the genotype G-14 is not significant and the regression co-efficient β_i is around unity (0.921) and as such it is regarded as stable variety. Similarly the deviation for genotype G-7 is not significant at 1% level and the coefficient of linear sensitivity is very close to unity, hence this can also be regarded as stable variety. The genotype G-6 tops in Table 2 with respect to the average yield over the environments. However the significance of deviations from linear regression makes its behaviour unpredictable over the environments and one may not be able to comment on its stability from Eberhart and Russell's model point of view.

AMMI for Balanced Groundnut Data

The AMMI model was evaluated for the balanced groundnut data. The corresponding AMMI analysis of interaction is presented in Table 4, which may elucidate the post-dictive success of AMMI model to explain the interaction.

Post-dictive Success

The first six PCA axes were found to be significant and hence retained in the model. The fitted AMMI model with first six PCA axes explained 89% of the interaction variation. The sum of squares accounted for by different significant axes and the residual are also presented in Table 4. The requirement of too many axes here indicates the complexity of G x E interaction pattern in the groundnut data. However, a large amount of non-linear interaction, Joint Regression failed to explain, is now explained by the AMMI model. Even with a single axis (PCA-1), the contribution is 40.3% as against the contribution of linear component of interaction in Joint Regression, 3.5%.

Predictive Accuracy

To estimate σ_V^2, we need atleast two replicates hence we were to construct the AMMI model for balanced groundnut data with a single replicate. One replicate out of three is chosen at random for each of the 300 treatments and

utilised for modelling. The AMMI model was fitted by retaining the first 6 PCA axes for the interaction. Now σ^2_{MV} is estimated empirically as the mean square difference between the model's estimates and the validation observations which turns out to be $\hat{\sigma}^2_{MV}$ = 130088.69. Likewise, σ^2_V is estimated empirically by the Error MS obtained from the one way ANOVA model, which turns out to be $\hat{\sigma}^2_V$ = 69719.5. Now the model's predictive accuracy may be assessed from the estimate of model's variance, $\hat{\sigma}^2_M$, which may be obtained as

$$\hat{\sigma}^2_M = \hat{\sigma}^2_{MV} - \hat{\sigma}^2_V = 60369.19$$

Now the effective number of replications may be evaluated as $\frac{\hat{\sigma}^2_V}{\hat{\sigma}^2_M} = 1.155$

Here the effective number of replications exceeded the actual replications supplied (one) to the model, hence the model exhibits the Stein effect. This means that the model's estimates are more predictively accurate, implying that the model is better than its data.

Biplots

The scaling constant c is chosen equal to 0.5 to obtain the genotypic and environmental scores.

Biplot with First PCA Axis

Figure 1 depicts the first PCA scores of genotypes and environments of balanced groundnut data plotted against their respective means. From Figure 1, the genotypes G-6, G-11, and G-12 may be considered as stable. Similarly the environments E-20, E-9, E-4 and E-7 may be regarded as stable. The genotypes G-4 and G-5 and the environment E-17 have larger interaction effects, hence may be regarded as unstable. There is huge variability in environmental means as compared to genotypic means. The genotypes G-4 and G-5 have negative interactions in E-17 and positive interaction in E-3. Similarly G-3 has positive interaction in E-17 but negative interaction in E-3. If we employ the strategy, simultaneous selection for yield and stability for the genotypes, G-6 is found to be superior to all others. However the stability conclusions drawn from this plot, which accounted for only 40.3% of interaction SS, are not precise.

Biplot with First Two PCA Axes

Figure 2 depicts the second PCA scores of genotypes and environments of balanced groundnut data plotted against their respective first PCA scores. From Figure 2, G-11 is the most stable among all the genotypes, which was at second rank in Figure 1. From Figure 2 also, G-4 and G-5 were found to be most unstable varieties. The genotype G-6, which was at first rank in Figure 1, now

goes to second rank. In Figure 2 also, E-17 was found to be the unstable environment. Interaction of a given genotype in a given environment may be obtained by projecting either vector on to the other, times the length of the vector on which projection takes place. If the given genotype and the environment are in the same quadrant their interaction will be positive.

The scope of biplots to efficiently explain the interaction is very much limited for the dataset considered here as is evident from the fact that the first PCA axis accounted for only 40.3% of the interaction variation and the first two PCA axes together constituted only 55% of the interaction variation. The postdiction suggested to retain more than two (six) PCA axes in the model to explain the interaction. In such cases it is not advisable to do stability conclusion from biplots. The stability measure that considers all the significant axes or all possible axes should be explored.

Measures of Stability from AMMI Model

The four measures of stability FP_i, B_i, FA_i and $W_{i(AMMI)}$ described earlier were evaluated for each of the 15 genotypes of balanced groundnut data. The 15 genotypes were ranked with respect to their stability with each of the four measures such that lesser the value of the rank, more is the stability. The stability rank orders displayed by these four measures of stability were presented in Table 5. Stability rank order displayed by FP_i in Table 5 reveals that genotypes G-6, G-11 and G-12 are the most stable genotypes in the descending order; and variety-4 is the least stable. The same stability conclusions were obtained from biplot in Figure 1. However the stability ranks of G-6 and G-11 are interchanged with B_i as compared to FP_i, the same conclusion obtained from biplot in Figure2. Only the ranks of genotypes G-3, G-4 and G-5 remained same to that of FP_i. The stability rank order displayed by FA_i (retaining 6 PCA axes in the AMMI model) in Table 5 is highly deviating from the ones obtained by FP_i as well as B_i; as we shall see the G-11 and G-6 are taking the ranks 6 and 8 respectively with FA_i. FA_i identified G-7 and G-5 as the most stable and least stable varieties respectively. The stability rank order displayed by $W_{i(AMMI)}$ in Table 5 reveals that G-14 is the most stable one and G-7 is the second most stable variety; just the ranks obtained by FA_i are interchanged. In contrast to FP_i and B_i, the G-5 was identified as the least stable variety by FA_i and $W_{i(AMMI)}$.

We may also observe in Table 5, the rank of G-14 is monotonically improving towards $W_{i(AMMI)}$. G-7 also exhibits almost a similar trend. This implies that there is some association improving towards the right ($W_{i(AMMI)}$). This trend supports the presumption made earlier that the reliability of the measure improves towards the right; with the increase in the number of PCA axes retained in the model. Now, we may use the tool Spearman's rank correlation to quantify the reliability of FP_i, B_i and FA_i assuming that the stability conclusions derived from $W_{i(AMMI)}$ as true. The rank correlations of stability rank orders displayed by FP_i, B_i and FA_i with rank order obtained by

$W_{i(AMMI)}$ were 0.761, 0.846 and 0.961 respectively. This increasing trend in rank correlations, also supports our presumption that the reliability of the measures are in the following order: $FP_i \leq B_i \leq FA_i \leq W_{i(AMMI)}$ i.e. the order in which amount of information increases.

AMMI for Incomplete Two-way Table

To evaluate the methodologies described to deal with missing data, unbalancedness is created by eliminating 20 cells at random in Table 1. The genotype, environment combinations identified for deletion are as following: (1,17), (2,3), (3,9), (4,1), (4,16), (5,5), (7,6), (8,5), (8,10), (9,7), (10,2), (10,15), (10,18), (11,11), (11,20), (12,4), (12,12), (13,6), (14,14) and (15,8). The three EM-AMMI approaches described for incomplete datasets were employed for the unbalanced groundnut data so derived. 50 iterations were found to be sufficient to achieve convergence in the imputed values with all the 3 approaches.

Predictive success of indirect data vs total data

$$\hat{\sigma}^2_{MV} = \hat{\sigma}^2_M + \hat{\sigma}^2_V$$

$$\Rightarrow \quad \hat{\sigma}^2_M = \hat{\sigma}^2_{MV} - \hat{\sigma}^2_V$$

If our aim is only the comparison of total and indirect data with respect to their predictive success $\hat{\sigma}^2_{MV}$ serves our purpose rather than $\hat{\sigma}^2_M$, provided validation data is same for both the models. When validation data is same, $\hat{\sigma}^2_V$ remains same for both the models and the two models may be compared in terms of $\hat{\sigma}^2_{MV}$ rather than $\hat{\sigma}^2_M$. $\hat{\sigma}^2_{MV}$ may directly be obtained as the square of RMSPD and we are not required to estimate σ^2_V. When we do not need to estimate σ^2_V, only one replicate for validation data suffice, which leaves (R – 1) replicates at our disposal for modelling data ensuring maximum possible precision for the model. As our groundnut data is having 3 replicates, we may take one replicate at random for each of the 300 treatments for validation data and the remaining two replicates (2 × 300 = 600 observations) for modelling data. Now for any treatment, the two replicates used for modelling constitute direct data and the indirect data consist of the remaining 299 treatments each having 2 replicates; so the indirect data contain 598 observations. Obviously the total data consist of both direct and indirect data *i.e.*, 2 + 598 = 600 observations.

AMMI model was fitted for the total data (600 observations) retaining first 6 PCA axes for the interaction. Root mean square predictive difference (RMSPD) between AMMI model's estimates of 300 treatments and their respective validation observations was found to be 313.89. The indirect data estimates are possible given an implementation of EM-AMMI. For each treatment in turn, EM-AMMI was given only for the indirect modelling data,

namely, the 600 – 2 = 598 other observations, and the model was used to impute a yield value for the missing cell. The result was a complete matrix of imputed yield values for all the 300 treatments. The RMSPD between these imputed values and their corresponding validation observations was evaluated which turned out to be 599.03.

Unfortunately, unlike the soyabean data of Gauch and Zobel [6], the groundnut data considered here for study does not exhibit a considerable amount of indirect EM-AMMI derived information to that of total information. The predictive success of indirect model is too poor as compared to the total model, for the data considered here.

Comparison of the Three EM-AMMI Approaches

The missing cells in the two-way table of incomplete groundnut data were imputed using EM-AMMI (Gauch and Zobel [6]) as well as Modified EM-AMMI (Bajpai [1]). Now the four measures of stability proposed were evaluated for the complete two way table so obtained. Stability rank order with each stability measure is obtained such that lesser the numerical value, more is the stability for the genotype. Table 6 and Table 7 correspond to the stability rank orders displayed by the four measures of stability with EM-AMMI (Gauch and Zobel [6]) and Modified EM-AMMI (Bajpai [1]). Table 6 and Table 7 also report the Spearman's rank correlations between rank order displayed by the measures of stability with incomplete data and the ones obtained with the corresponding complete data.

In case of EM-AMMI with random environments, the results need to be compared with a similar approach applied to the complete data. The two-way table of interactions are obtained by correcting balanced data for the BLUE's of μ and α_i and the BLUP's of β_j obtained from mixed model equations. Now, the four measures of stability are evaluated for the two-way table of interactions so obtained and the corresponding stability rank orders are presented in Table 8.

The stability rank orders obtained using the approach EM-AMMI with random environments for incomplete groundnut data are presented in Table 9. Table 9 also reports the rank correlations between the stability rank order obtained with incomplete groundnut data using EM-AMMI with random environments and the ones obtained with complete (balanced) groundnut data using AMMI with random environments, presented in Table 8.

One may be interested to compare Table 5 with Table 8 to see what happened when the environments are treated as random. This comparison reveals that the stability rank order remained same for all the stability measures from Table 5 to Table 8 except for $W_{i(AMMI)}$ with which the ranks of genotypes G-1 and G-8 got interchanged. The results in Table 8 may be preferred to Table 5; since treating environmental effects as random has got a desirable implication of precluding bias creeping in, due to selection of environments.

Now two questions will arise; which stability measure out of the four is to rely upon and which EM-AMMI approach out of the three is to be employed in case of incomplete data. If we compare the rank correlations under FP_i, B_i, FA_i and $W_{i(AMMI)}$ in the three EM-AMMI approaches it may lead to the following conclusions. In all the three EM-AMMI approaches the rank correlations under FP_i, B_i and FA_i are found to be poor (less than or equal) as compared to $W_{i(AMMI)}$ which implies that the stability measures FP_i, B_i and FA_i are more sensitive to missing of data and $W_{i(AMMI)}$ exhibits some kind of robustness to the missing of data. Hence we may conclude that $W_{i(AMMI)}$ is not only reliable as compared to the remaining 3 measures, but also robust to the missing observations. The rank correlations under the measures of stability in EM-AMMI with random environments are found to be better than or equal to the ones obtained with the other two approaches except for FA_i which was found to be better in EM-AMMI (Gauch and Zobel [6]). As we are more interested in $W_{i(AMMI)}$ on account of its high reliability and robustness, we may compare the rank correlations under $W_{i(AMMI)}$ for the three EM-AMMI approaches which reveals that EM-AMMI with random environments is superior to the rest two having the stability rank order more closer to the one displayed by balanced data. Besides, the approach EM-AMMI with random environments has some theoretically desirable consequences, in the sense that it precludes the bias creeping in due to the selection of environments. These considerations may motivate us to prefer and recommend $W_{i(AMMI)}$ using EM-AMMI with random environments for incomplete data sets over the other two approaches to rank the genotypes with respect to their stability.

EM-AMMI Enriched Technique for Joint Regression for Unbalanced Data

The estimated variety parameters with EM-AMMI Substituted least squares and EM-AMMI Substituted BLUP for Joint Regression are presented in Table 10. We may observe in Table 10, the estimates of variety parameters obtained from EM-AMMI Substituted least squares are very close to that of EM-AMMI Substituted BLUP. We may compare the results of Table 10 with the results of Table 2, which reveals that employment of EM-AMMI enriched technique for Joint Regression for unbalanced data yields almost the same conclusions to that of balanced data.

REFERENCES

[1] Bajpai, P.K. (1998). Some contributions to improved estimation of genotype environment interaction and analysis of genotypic stability with special reference to sugarcane crop. Ph.D. Thesis, P.G. School, IARI, New Delhi.

[2] Eberhart, S.A. and Russel, W.A. (1966). Stability parameters for comparing varieties. *Crop Science*, **6**, 36-40.

[3] Gauch, H.G. (1988). Model selection and validation for yield trials with interaction. *Biometrics*, **44**, 705-715.

[4] Gauch, H.G. (1992). Statistical analysis of regional yield trials. *Elsevier*, Amsterdam.

[5] Gauch, H.G. and Zobel, R.W. (1988). Predictive and postdictive success of statistical analyses of yield trials. *Theoretical and Applied Genetics*, **76**, 1-10.

[6] Gauch, H.G. and Zobel, R.W. (1990). Imputing missing yield trial data. *Theoretical and Applied Genetics*, **79**, 753-761.

[7] Gollob, H.F. (1968). A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika*, **33**, 73-115.

[8] Little, R.J.A. and Rubin, D.B. (1987). *Statistical Analysis with Missing Data*. Wiley and Sons, New York.

[9] Shukla, G.K. (1972). Some statistical aspects of partitioning genotype environmental components of variability. *Heredity*, **29**, 237-245.

[10] Stein, C. (1955). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In: *Proc 3rd Berkeley Symp Math Stat Prob*, University of California Press, Berkeley CA, 197-206.

[11] Wricke, G. (1965). Die Erfassung der Wechselwirkungen zwischen genotyp und umwelt bei quantitativen Eggenschaften. *Zeitschrift fur pflanzenzuchtung*, **53**, 266-343.

APPENDIX

Table 1. Mean data of groundnut

	E-1	E-2	E-3	E-4	E-5	E-6	E-7	E-8	E-9	E-10	E-11	E-12	E-13	E-14	E-15	E-16	E-17	E-18	E-19	E-20	Mean
G-1	1773	880	2841	2020	856	1382	1458	282	1190	1001	2708	1832	1188	2252	1583	2014	2199	810	1033	992	1515
G-2	1715	861	2497	2020	505	1104	1153	275	1394	882	1956	1907	729	1658	1285	1986	2014	865	600	842	1312
G-3	1241	424	3266	1717	1148	1225	1130	113	701	705	1688	1568	1153	2073	1303	2361	2893	1028	1000	997	1387
G-4	1472	917	3172	2222	1505	1475	1222	632	1308	334	2833	1157	792	956	1374	2570	611	486	333	1049	1321
G-5	1208	1435	3625	1919	903	1432	921	862	1081	539	2303	1778	577	1132	1368	2691	495	639	300	877	1304
G-6	1893	1310	2716	2374	1320	1476	1482	680	1468	591	2877	2333	1005	2636	1438	2812	1968	963	1100	1413	1693
G-7	1852	1169	2527	2222	903	1220	1407	455	1637	521	2042	1732	1285	2046	1333	2500	2060	949	633	877	1469
G-8	1266	993	2245	1869	292	972	1171	275	1419	767	2184	2037	799	1749	1368	2083	1537	732	667	965	1270
G-9	1736	792	2376	2172	981	1113	1051	364	1579	364	2940	1500	819	1668	1041	1944	2431	1000	633	967	1374
G-10	1442	695	2800	2071	1051	1890	1051	605	1684	67	2083	1419	1146	1295	1750	2726	1713	50	600	1166	1365
G-11	1530	1055	2643	2172	1412	1049	1051	567	1211	174	1977	1963	1083	2063	1319	1789	1435	944	633	1309	1369
G-12	1697	1222	2770	2273	1759	1343	1153	572	1169	353	2014	2222	792	1634	1319	2271	2014	1176	1200	1026	1499
G-13	1637	1097	2715	2071	1806	1158	1199	636	1269	437	1574	1843	958	1719	1299	2014	2431	1014	1033	1379	1464
G-14	1641	1403	2712	2071	792	1037	1199	757	1296	643	2347	1889	1035	1551	1375	1993	2222	875	933	1092	1443
G-15	1727	1139	2452	2071	481	883	1519	299	1330	366	1535	1574	1070	1940	1146	1514	2208	745	567	904	1274
Mean	1589	1026	2757	2084	1048	1251	1211	492	1316	516	2204	1784	962	1758	1353	2218	1882	818	751	1057	1404

Table 2. Stability parameters for the 15 genotypes of groundnut

Genotype	$\hat{a}_i$	$\hat{\beta}_i$	$\overline{S}^2_{d_i}$
G-1	1514.75	1.034	38312
G-2	1312.4	0.961	20484
G-3	1386.78	1.096	145570
G-4	1320.95	1.100	191293
G-5	1304.27	1.110	207790
G-6	1692.65	1.109	31631
G-7	1468.53	0.998	15067
G-8	1269.47	0.918	36896
G-9	1373.57	1.074	63579
G-10	1365.13	1.095	106905
G-11	1369.05	0.910	32370
G-12	1498.95	0.942	39707
G-13	1464.43	0.833	52999
G-14	1443.25	0.921	10586
G-15	1273.48	0.899	66874

Table 3. Analysis of variance for the balanced groundnut data with eberhart and russel model

Source	df	MS
Genotypes	14	254686 **
Env + Gen × Env	285	
Env (linear)	1	
Gen × Env (linear)	14	62925 NS
Pooled deviation	270	90839
G-1	18	58480 **
G-2	18	40652 **
G-3	18	165738 **
G-4	18	211461 **
G-5	18	227958 **
G-6	18	51799 **
G-7	18	35235 *
G-8	18	57064 **
G-9	18	83747 **
G-10	18	127073 **
G-11	18	52538 **
G-12	18	59875 **
G-13	18	73167 **
G-14	18	30754 NS
G-15	18	87042 **
Average error (pooled error)	560	20168

Table 4. AMMI analysis of variance for the interaction of balanced groundnut data

Source	df	SS	MS	Variance ratio
G × E interaction	266	25408292	95519.89	
PCA –1	32	10240492	320015.38	12.05
PCA –2	30	3899700.4	129990.01	4.90
PCA –3	28	2795398.8	99835.67	3.76
PCA –4	26	2377000.8	91423.11	3.44
PCA –5	24	1961588.9	81732.87	3.08
PCA –6	22	1372729.6	62396.80	2.35
Residual	104	2761382.0	26551.75	

Table 5. Stability rank orders displayed by FP_i, B_i, FA_i and $W_{i(AMMI)}$with balanced groundnut data

Genotype	FP_i	B_i	FA_i	$W_{i(AMMI)}$
G-1	9	6	5	6
G-2	8	7	3	3
G-3	13	13	13	13
G-4	15	15	14	14
G-5	14	14	15	15
G-6	1	2	8	5
G-7	6	4	1	2
G-8	4	9	7	7
G-9	7	5	11	10
G-10	11	10	12	12
G-11	2	1	4	4
G-12	3	8	6	8
G-13	10	12	10	9
G-14	5	3	2	1
G-15	12	11	9	11
Rank Correlations	0.761	0.846	0.961	1

Table 6. Stability rank orders displayed by the four measures of stability with incomplete groundnut data using EM-AMMI (Gauch and Zobel [6])

Genotype	FP_i	B_i	FA_i	$W_{i(AMMI)}$
G-1	12	12	10	10
G-2	4	4	4	4
G-3	13	13	13	13
G-4	14	14	15	15
G-5	15	15	14	14
G-6	2	1	7	7
G-7	5	5	3	2
G-8	3	2	2	3

Genotype	FP_i	B_i	FA_i	$W_{i(AMMI)}$
G-9	6	11	8	8
G-10	10	10	12	12
G-11	8	8	6	5
G-12	1	3	5	6
G-13	9	7	9	9
G-14	7	6	1	1
G-15	11	9	11	11
Rank Correlations	0.8607	0.5607	0.8607	0.9143

Table 7. Stability rank orders displayed by the four measures of stability with incomplete groundnut data using Modified EM-AMMI (Bajpai [1])

Genotype	FP_i	B_i	FA_i	$W_{i(AMMI)}$
G-1	12	12	10	10
G-2	5	2	4	4
G-3	13	13	13	13
G-4	14	14	14	14
G-5	15	15	15	15
G-6	1	5	7	7
G-7	4	1	3	2
G-8	3	3	1	3
G-9	7	10	8	8
G-10	10	8	12	12
G-11	6	7	6	6
G-12	2	4	5	5
G-13	9	9	9	9
G-14	8	6	2	1
G-15	11	11	11	11
Rank Correlations	0.9036	0.6143	0.8464	0.9036

Table 8. Stability rank orders displayed by FP_i, B_i, FA_i and $W_{i(AMMI)}$ with balanced groundnut data using AMMI with random environments

Genotype	FP_i	B_i	FA_i	$W_{i(AMMI)}$
G-1	9	6	5	7
G-2	8	7	3	3
G-3	13	13	13	13
G-4	15	15	14	14
G-5	14	14	15	15
G-6	1	2	8	5
G-7	6	4	1	2
G-8	4	9	7	6
G-9	7	5	11	10
G-10	11	10	12	12

Genotype	FP_i	B_i	FA_i	$W_{i(AMMI)}$
G-11	2	1	4	4
G-12	3	8	6	8
G-13	10	12	10	9
G-14	5	3	2	1
G-15	12	11	9	11

Table 9. Stability rank orders displayed by FP_i, B_i, FA_i and $W_{i(AMMI)}$ with incomplete groundnut data using EM- AMMI with random environments

Genotype	FP_i	B_i	FA_i	$W_{i(AMMI)}$
G-1	12	12	10	10
G-2	5	1	4	4
G-3	13	13	13	13
G-4	14	14	14	14
G-5	15	15	15	15
G-6	1	4	7	7
G-7	4	2	3	2
G-8	3	3	1	3
G-9	7	10	8	9
G-10	10	8	12	12
G-11	6	7	6	6
G-12	2	5	5	5
G-13	9	9	9	8
G-14	8	6	2	1
G-15	11	11	11	11
Rank Correlations	0.9036	0.625	0.8464	0.9321

Table 10. Estimated variety parameters with EM-AMMI substituted least squares and EM-AMMI substituted BLUP for joint regression

Genotype	EM-AMMI substituted least squares		EM-AMMI substituted BLUP	
	α_i	β_i	α_i	β_i
G-1	1542.54	1.071	1542.56	1.084
G-2	1335.43	1.039	1335.46	1.052
G-3	1389.17	1.096	1389.07	1.109
G-4	1273.24	0.995	1272.37	1.005
G-5	1306.02	1.093	1305.44	1.106
G-6	1692.75	1.104	1692.75	1.117
G-7	1477.35	0.988	1477.16	1.000
G-8	1273.83	0.962	1273.91	0.973
G-9	1391.03	1.072	1391.02	1.085
G-10	1395.47	1.000	1396.04	1.012
G-11	1391.94	0.999	1390.94	1.009
G-12	1471.53	0.907	1471.50	0.918
G-13	1468.06	0.834	1468.20	0.845
G-14	1476.30	0.953	1476.34	0.965
G-15	1278.40	0.887	1278.40	0.898

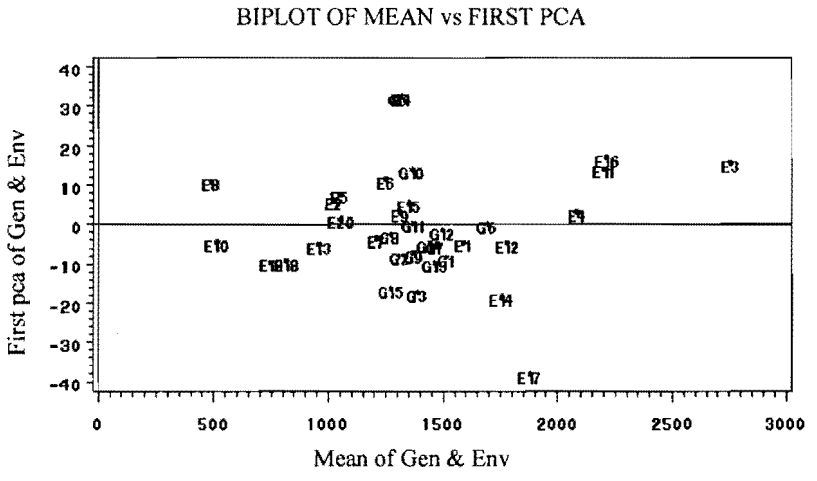

Fig. 1. Biplot of means Vs first PCA for the balanced groundnut data

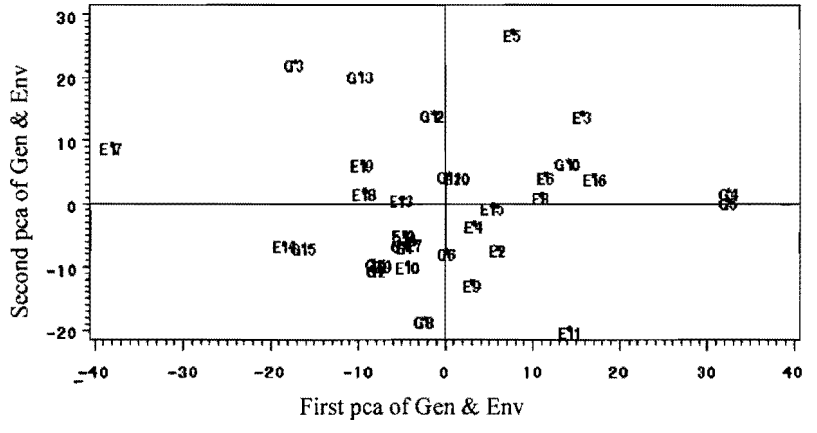

Fig. 2. Biplot of first PCA Vs second PCA for the balanced groundnut data